

**REMARKS**

The amendments to claim 53 and new claims 65 and 66 are supported by the specification at page 16, lines 16-20 and 26-30, page 17, lines 12-13, and page 33, lines 26-33.

**Claim Rejections – 35 USC § 112, First Paragraph**

The Examiner rejected the claims for lack of written description (possession) for the term CTBF. In the present response, the claims are now modified to recite that the CTBF is an organic molecule less than 2000 daltons and comprises a homocyclic aromatic radical, a heterocyclic aromatic radical or a heterocyclic radical. In the present specification, Figures 2 and 4 also provide 41 specific examples of CTBF radicals. The specification states that “[c]andidate target binding fragments that find use in the herein described method may be employed in the herein described method as originally obtained from a commercial or noncommercial source (for example, a large number of small organic chemical compounds that serve as candidate target binding fragments are readily obtainable from commercial suppliers such as Aldrich Chemical Co., Milwaukee and Sigma Chemical Co., St. Louis, MO) or maybe obtained from chemical synthesis.”

Applicants contend that present invention meets the statutory requirements for written description, given the description provided in the specification and the fact that libraries of molecules for use in pharmacological assays were known in the art as of the filing date. A rejection for lack of written description for the present claims is improper. See MPEP Section 2163, III A. (B), which states “A general allegation of ‘unpredictability in the art’ is not a sufficient reason to support a rejection for lack of adequate written description.” (Emphasis added.) See also MPEP Section 2163.03, which states:

While a question as to whether a specification provides an adequate written description may arise in the context of an original claim which is not described

sufficiently (see, e.g., *Regents of the University of California v. Eli Lilly*, 119 F.3d 1559, 43 USPQ2d 1398 (Fed. Cir. 1997)), there is a strong presumption that an adequate written description of the claimed invention is present in the specification as filed. *In re Wertheim*, 541 F.2d 257, 262, 191 USPQ 90, 96 (CCPA 1976). Consequently, rejection of an original claim for lack of written description should be rare. (Emphasis added.)

Thus, it is incumbent upon the Examiner to present reasons of lack of written description that will overcome a strong presumption of adequate written description.

How is written description shown? As MPEP § 2163.03 discusses, written description is shown by whether the disclosure of the application "reasonably conveys to the artisan that the inventor had possession at that time of the later claimed subject matter." *Ralston Purina Co. v. Far-Mar-Co., Inc.*, 772 F.2d 1570, 1575, 227 USPQ 177, 179 (Fed. Cir. 1985) (quoting *In re Kaslow*, 707 F.2d 1366, 1375, 217 USPQ 1089, 1096 (Fed. Cir. 1983)).

This begs the question, how is possession shown? As discussed in MPEP § 2163.03 "[a]n applicant shows possession of the claimed invention by describing the claimed invention with all of its limitations using such descriptive means as words, structures, figures, diagrams, and formulas that fully set forth the claimed invention." *Lockwood v. American Airlines, Inc.*, 107 F.3d 1565, 1572, 41 USPQ2d 1961, 1966 (Fed. Cir. 1997). (Emphasis added.) This same section then elaborates "description of an actual reduction to practice, or by showing that the invention was "ready for patenting" such as by the disclosure of drawings or structural chemical formulas that show that the invention was complete, or by describing distinguishing identifying characteristics sufficient to show that the applicant was in possession of the claimed invention." Applicants have described the CTBF as an organic molecule less than 2000 daltons and comprises a homocyclic aromatic radical, a heterocyclic aromatic radical or a heterocyclic radical. As of the earliest priority date of the application, these terms were well understood by one of skill in the art and therefore fulfill the statutory requirement of written description. Figures 2 and 4, as well as the Aldrich and Sigma catalogues, demonstrate that the applicants had possession of a genus of small organic molecules.

To support the argument of lack of written description, the Examiner relies on *Lilly*. (The cases cited by the Examiner are related to enablement and utility, which are separate statutory requirements.) The facts of *Lilly* are clearly distinguishable from the present case. *Lilly* was a biotechnology case related to cDNA sequences. Before the time of the *Lilly* invention, the cDNA sequences that encoded insulin were unknown. The inventors of *Lilly* sequenced the cDNA for rat insulin. Based on this one described species, the claims at issue in *Lilly* were drawn to cDNA sequences for all mammalian insulin. None of the structures of the other cDNA sequences were known, nor could they be derived from the rat cDNA sequence. The Federal Circuit accordingly held that in this case, a single known cDNA species did not provide written description of any of the other cDNA species (such as human insulin), or a genus consisting of only one described cDNA species and with all of the other cDNA species not described.

In contrast to these facts of *Lilly*, at the time to the filing of the application, numerous chemical structures for species of the genus CTBFs have been known for decades. One of skill in the art, when reading these terms, in light of Figures 2 and 4 of the specification, the Aldrich and Sigma catalogs and libraries known in the art would be able to immediately envision numerous examples of CTBFs. Unlike *Lilly*, the CTBF is not a single species that is being used to describe an entire genus of unknown chemical species. Therefore, the application of *Lilly* to the present claims is improper and the rejection should be withdrawn.

#### **Claim Rejections – 35 USC § 112, Second Paragraph**

The rejections for indefiniteness have been overcome as follows. Claim 53 has been amended to recite that the two CTBF's are combined together in a microtiter well. The present claims no longer contain the term small. Finally, claims referencing a second LG have been deleted.

**Claim Rejections – 35 USC § 103(a)**

The present rejection for obviousness combines two references: Kirkpatrick and Silverman. Applicants contend that the Examiner has not established a *prima facie* case of obviousness, which requires (1) a teaching or suggestion of all the elements of the claims in the cited references, (2) motivation to combine or modify references to arrive at the present invention, and (3) a reasonable expectation of success of the invention. See MPEP § 706.02(j). Applicants contend the present rejection for obviousness fails to meet all three requirements.

First, as previously admitted by the Examiner in, *inter alia*, at page 5 of the Advisory Action dated November 17, 2004, “Kirkpatrick et al., do teach a ‘mixture’ of library members.” The Examiner has not argued that Silverman teaches a combinatorial library, much less a combinatorial library where the library members are combined together in a microtiter well.

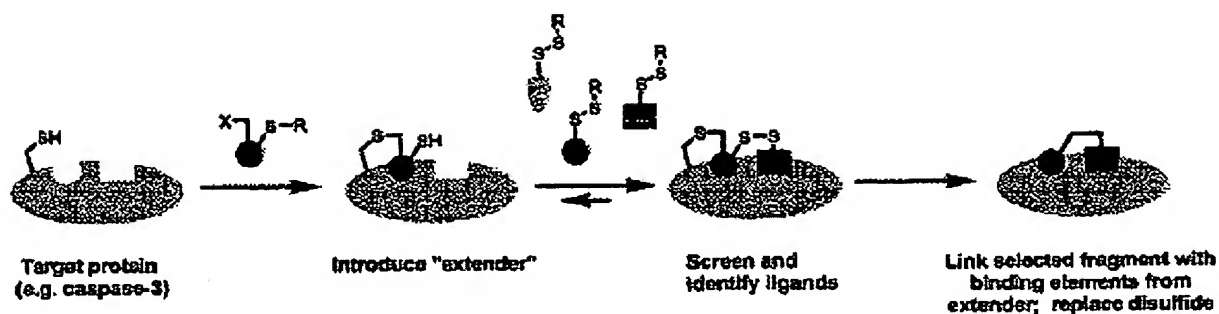
Table 2.2 of Silverman is cited by the Examiner to support the proposition that in biochemistry a  $-Cl$  radical is equivalent to  $-NH_3$  radical. Applicants respectfully request reconsideration of this proposition because this is a misreading of the Silverman reference. Table 2.3 of Silverman, actually relates to “classical isoterres” and not bioisoterres. According to Silverman at page 19, classical isoterres were defined by Grimm in 1925 as “groups that have the same number of valence electrons but may have a different number of atoms.” Erlenmeyer later defined isoterres as “atoms, ions or molecules in which the peripheral layers of electrons can be considered identical.” Therefore, what are termed isoterres, are atoms that are only related by the number of valence electrons. One of skill in the art, indeed one who has completed an introductory college level chemistry course, would clearly understand that merely because atoms have the same number of valence electrons, they would not have significantly similar chemical properties. Indeed, Table 2.2 of Silverman proffers that carbon and silicon are isoterres, but one of skill in the art would not entertain that silicon atoms are obvious substitutes for or even similar to carbon atoms as studied in the field of organic chemistry (i.e. essentially the chemistry of hydrocarbons) and biochemistry. Likewise, the

fact that  $-\text{CH}_3$ ,  $-\text{NH}_2$ ,  $-\text{OH}$ ,  $-\text{F}$  and  $-\text{Cl}$  are listed as “classical isoterres”, does not mean that these molecules are considered by Silverman to be bioisoterres.

One of skill in the art would recognize that an alkyl substituted with  $-\text{CH}_3$ ,  $-\text{NH}_2$ ,  $-\text{OH}$ ,  $-\text{F}$  or  $-\text{Cl}$  would have different chemical properties. It is noteworthy that alkyl chlorines are insoluble in aqueous solution and are notorious carcinogens. Alkyl amines, however, are able to form hydrogen bonds with water and are, therefore, considered to soluble in aqueous solutions and are not generally considered carcinogens.

Moreover, there is no motivation or reasonable expectation of success in combining Silverman with Kirkpatrick. If anything, Silverman stands for the proposition that the “nonclassical bioisoterres” of Table 2.3, can be interchanged for “lead modification”. (See page 21.) (Emphasis added.) Lead modification would be relevant to modifying the CTBF, not the  $-\text{S}-\text{S}-\text{R}^8$  portion of the library. Moreover, in Table 2.3, in which related “nonclassical bioisoterres”, halogens (item 5) are described as being bioisoterres of  $-\text{CF}_3$ ,  $-\text{CN}$ ,  $-\text{C}(\text{CN})_2$ , or  $-\text{C}(\text{CN})_3$ . A  $-\text{NH}_2$  is not described as being a bioisotere of  $-\text{Cl}$ . Silverman defines on page 21 bioisoterres as molecules that “do not have the same number of atoms and do not fit steric and electronic rules of classical isoterres, but they do produce a similarity in biological activity.”

Moreover, as the attached article by Erlanson, notes, an amine moiety is useful in a biological assay because it improves solubility. Erlanson also describes how an approach to combinatorial chemistry called extended tethering combines the benefits of dynamic combinatorial chemistry with covalent tethering for molecules that have a cysteine residue near the active site. Such a residue can be naturally occurring or one created by mutagenesis. The steps involved with using the present invention for extended tethering are shown below. The library of the present invention is represented by the CTBF-S-S-R compounds below, wherein R is the same as  $\text{R}^8$  of amended claim 1.



In the first step of using the library of the present invention, a cysteine residue in or near the active site is modified with a extender molecule, which is a molecule of some inherent affinity for the protein of interest that bonds to the cysteine and also contains a protected thiol group. After deprotection, the protein complexed with the extender is screened against a library of claim 1. CTBFs that have affinity, even if it is weak, will form stable disulfide bonds with the extender. The complexes are then sufficiently stable to be detected by mass spectrometry. Detection of such weakly binding complexes substantially facilitates the process of lead optimization.

Without impressive hindsight based on the teachings of the present specification one of ordinary skill in the art would not have been motivated to combine Kirkpatrick and Silverman. Accordingly, applicants urge that the present claims to be in condition for allowance. The Examiner is invited to contact the undersigned by telephone to advance prosecution of this application.

**Conclusion**

The Commissioner is hereby authorized to charge any additional fees which may be required regarding this application under 37 C.F.R. §§ 1.16-1.17, or credit any overpayment, to Deposit Account No. 19-0741. Should no proper payment be enclosed herewith, as by a check being in the wrong amount, unsigned, post-dated, otherwise improper or informal or even entirely missing, the Commissioner is authorized to charge the unpaid amount to Deposit Account No. 19-0741.

Respectfully submitted,

August 19, 2005

Date



Matthew E. Mulkeen  
Attorney for Applicants  
Registration No. 44,250

FOLEY & LARDNER LLP  
Customer Number: 22428  
Telephone: (202) 672-5446  
Facsimile: (202) 672-5399

# Site-directed ligand discovery

Daniel A. Erlanson<sup>\*†</sup>, Andrew C. Braisted<sup>\*</sup>, Darren R. Raphael<sup>\*</sup>, Mike Randal<sup>\*</sup>, Robert M. Stroud<sup>‡</sup>, Eric M. Gordon<sup>\*</sup>, and James A. Wells<sup>\*†</sup>

<sup>\*</sup>Sunesis Pharmaceuticals, Incorporated, 3696 Haven Avenue, Suite C, Redwood City, CA 94063; and <sup>‡</sup>Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143-0448

Contributed by James A. Wells, June 12, 2000

We report a strategy (called "tethering") to discover low molecular weight ligands ( $\approx 250$  Da) that bind weakly to targeted sites on proteins through an intermediary disulfide tether. A native or engineered cysteine in a protein is allowed to react reversibly with a small library of disulfide-containing molecules ( $\approx 1,200$  compounds) at concentrations typically used in drug screening (10 to 200  $\mu$ M). The cysteine-captured ligands, which are readily identified by MS, are among the most stable complexes, even though in the absence of the covalent tether the ligands may bind very weakly. This method was applied to generate a potent inhibitor for thymidylate synthase, an essential enzyme in pyrimidine metabolism with therapeutic applications in cancer and infectious diseases. The affinity of the untethered ligand ( $K_i \approx 1$  mM) was improved 3,000-fold by synthesis of a small set of analogs with the aid of crystallographic structures of the tethered complex. Such site-directed ligand discovery allows one to nucleate drug design from a spatially targeted lead fragment.

The drug discovery process usually begins with massive screening of compound libraries (typically hundreds of thousands of members) to identify modest affinity leads ( $K_d \approx 1$  to 10  $\mu$ M). Although some targets are well suited for this screening process, most are problematic because moderate affinity leads are difficult to obtain. Identifying and subsequently optimizing weaker binding compounds would improve the success rate, but screening at high concentrations is generally impractical because of compound insolubility and assay artifacts. Moreover, the screening process does not target specific sites for drug design, only those sites for which a high-throughput assay is available. Finally, many traditional screening methods rely on inhibition assays that are often subject to artifacts caused by reactive chemical species or denaturants.

We have developed an alternative strategy to rapidly and reliably identify small soluble drug fragments (molecular weight  $\approx 250$  Da) that bind with low affinity to a specifically targeted site on a protein or macromolecule. This method relies on the formation of a disulfide bond between the ligand and a cysteine residue in the protein of interest (Fig. 1A). A library of disulfide-containing molecules (Fig. 1B) is allowed to react with a cysteine-containing target protein under partially reducing conditions that promote rapid thiol exchange. Most of these library members will show no intrinsic affinity for the protein, and therefore the associated disulfide bond to the protein will be easily reduced. However, if a molecule has even weak inherent affinity for the target protein, the disulfide bond will be entropically stabilized and the equilibrium will lie toward the modified protein. Tethered compounds can then be identified by MS. Furthermore, the tethered complex is amenable to analysis by x-ray crystallography, which greatly facilitates the optimization of affinity once the disulfide tether is removed.

As an initial test of this technology, we chose the enzyme thymidylate synthase (TS). TS is a critical component of the sole *de novo* pathway for synthesis of dTMP from dUMP (1, 2). Thus TS is a good anticancer and anti-infective drug target (3–6). TS has also previously been used as a model system with which to validate other methods of drug discovery (7–10).

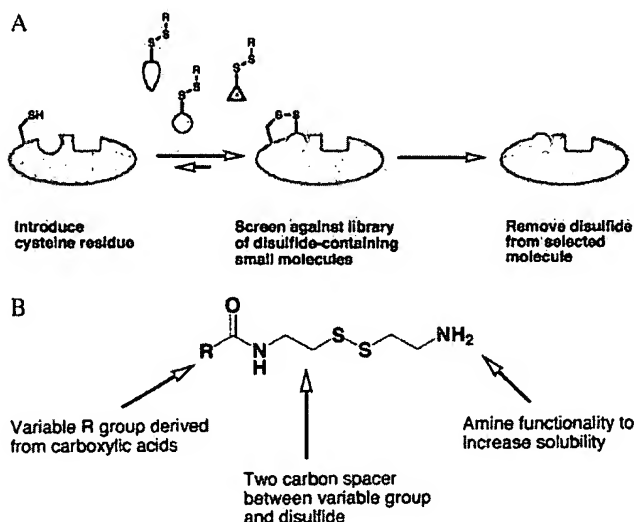


Fig. 1. (A) Schematic illustration of the tethering approach: a cysteine-containing protein is equilibrated with a disulfide-containing library in the presence of a reducing agent such as 2-mercaptoethanol. Most of the library members will have little or no inherent affinity for the protein, and thus by mass action the equilibrium will lie toward the unmodified protein. However, if a library member does show inherent affinity for the protein, the equilibrium will shift toward the modified protein. (B) Schematic illustration of a generic disulfide library derived from carboxylic acids. Other functional groups have also been converted to disulfide libraries, as described in *Materials and Methods*. In the present case, 1,200 compounds were screened against TS in pools of 8 to 15 compounds.

## Materials and Methods

**Thymidylate Synthase.** Using the standard techniques of molecular biology, we have created three mutants of the unmodified or "wild-type" *Escherichia coli* TS enzyme, overexpressed them in *E. coli* strain  $\chi$ 2913 [in which the TS gene has been eliminated (11)], and purified them. The  $\chi$ 2913 strain requires a thymidine supplement because the (deleted) TS gene is essential for life. We have constructed mutants including one in which the active-site cysteine has been replaced by serine (abbreviated as C146S) and two in which, in addition to this change, a new cysteine residue has been introduced into the active site, denoted C146S/L143C or C146S/H147C. All proteins have been expressed, purified, and characterized by SDS/PAGE and MS. Enzymatic inhibition assays were performed largely as described previously (12, 13).

Abbreviations: TS, thymidylate synthase;  $K_i$ , inhibition constant; BOC, butoxycarbonyl; DCM, dichloromethane; TFA, trifluoroacetic acid.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, [www.rcsb.org](http://www.rcsb.org) (PDB ID codes 1F4B, 1F4C, 1F4D, 1F4E, 1F4F, and 1F4G).

<sup>†</sup>To whom reprint requests should be addressed. E-mail: [jaw@sunesis-pharma.com](mailto:jaw@sunesis-pharma.com) or [erlanson@sunesis-pharma.com](mailto:erlanson@sunesis-pharma.com).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



**Reagents.** All commercially available materials were used as received. All synthesized compounds were characterized by  $^1\text{H}$  NMR [Bruker (Billerica, MA) DMX400 MHz Spectrometer] and HPLC-MS (Hewlett-Packard Series 1100 MSD).

**Disulfide Libraries.** The disulfide-containing library members were made from commercially available carboxylic acids and mono-*N*-(*tert*-butoxycarbonyl)-protected cystamine (mono-BOC-cystamine) by adapting the method of Parlow and coworkers (14). Briefly, 260  $\mu\text{mol}$  of each carboxylic acid was immobilized onto 130  $\mu\text{mol}$  equivalents of 4-hydroxy-3-nitrobenzophenone on polystyrene resin using 1,3-diisopropylcarbodiimide (DIC) in *N,N*-dimethylformamide (DMF). After 4 h at room temperature, the resin was rinsed with DMF ( $\times 2$ ), dichloromethane (DCM,  $\times 3$ ), and tetrahydrofuran (THF,  $\times 1$ ) to remove uncoupled acid and DIC. The acids were cleaved from the resin via amide formation with 66  $\mu\text{mol}$  of mono-BOC protected cystamine in THF. After reaction for 12 h at ambient temperature, the solvent was evaporated, and the BOC group was removed from the uncoupled half of each disulfide by using 80% trifluoroacetic acid (TFA) in DCM. The products were characterized by HPLC-MS, and those products that were substantially pure were used without further purification. After the initial discovery of the *N*-tosyl-D-proline lead (see below), a small number of other proline derivatives and *N*-tosyl-amino acids were added to the library. A total of 530 compounds were made by using this methodology.

Libraries (not shown) were also constructed from mono-BOC-protected cystamine and a variety of sulfonyl chlorides, isocyanates, and isothiocyanates. In the case of the sulfonyl chlorides, 10  $\mu\text{mol}$  of each sulfonyl chloride was coupled with 10.5  $\mu\text{mol}$  of mono-BOC-protected cystamine in THF (with 2% diisopropyl ethyl amine) in the presence of 15 mg of poly(4-vinylpyridine). After 48 h, the poly(4-vinylpyridine) was removed via filtration, and the solvent was evaporated. The BOC group was removed by using 50% TFA in DCM. In the case of the iso(thio)cyanates, 10  $\mu\text{mol}$  of each isocyanate or isothiocyanate was coupled with 10.5  $\mu\text{mol}$  of mono-BOC-protected cystamine in THF. After reaction for 12 h at ambient temperature, the solvent was evaporated, and the BOC group was removed by using 50% TFA in DCM. A total of 212 compounds were made by using this methodology.

Finally, oxime-based libraries were constructed by reacting 10  $\mu\text{mol}$  of specific aldehydes or ketones with 10.5  $\mu\text{mol}$  of  $\text{HO}(\text{CH}_2)_2\text{S-S}(\text{CH}_2)_2\text{ONH}_2$  (to be described separately) in 1:1 methanol/chloroform (with 2% acetic acid added) for 12 h at ambient temperature to yield the oxime product. A total of 448 compounds were made by using this methodology.

Individual library members were redissolved in either acetonitrile or dimethyl sulfoxide to a final concentration of 50 or 100 mM. Aliquots of each of these were then pooled into groups of 8–15 discrete compounds, with each member of the pool having a unique molecular weight.

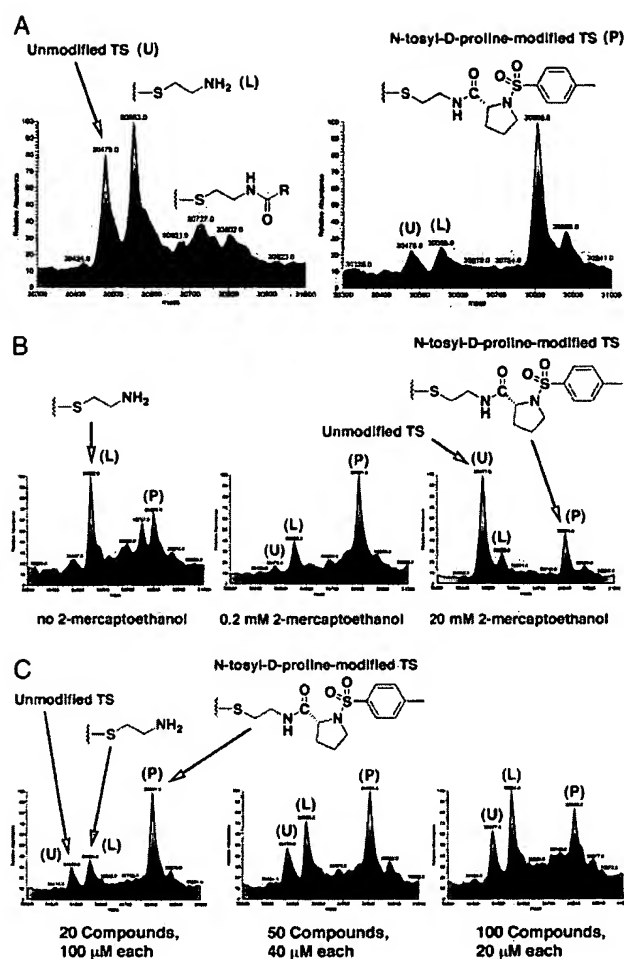
***N*-tosyl-D-proline Derivatives.** Synthesis of the *N*-tosyl-D-proline analogs began by reacting proline methyl ester hydrochloride with 4-(chlorosulfonyl)benzoic acid and sodium carbonate in water. The product was converted to the pentafluorophenyl ester by reacting it with pentafluorophenyl trifluoroacetate and pyridine in *N,N*-dimethylformamide and was purified via flash chromatography. This activated ester was then reacted with the methyl ester of glutamate (or any of the other amino acids tested) in the presence of triethylamine and dichloromethane, the product purified by flash chromatography and the methyl esters hydrolyzed with lithium hydroxide in water. The final products were purified via reverse-phase HPLC and lyophilized.

Alternatively, the above sequence was followed starting with proline *t*-butyl ester. After coupling of the amino ester to the benzoic acid, the *t*-butyl ester was removed with 50% TFA in DCM with triethylsilane as a scavenger. The free acid was then converted to a pentafluorophenyl ester as above and reacted with the appropriate amine. The methyl esters were hydrolyzed with lithium hydroxide in water, and the final products were purified via reverse-phase HPLC and lyophilized. Full synthetic details will be presented elsewhere.

**Disulfide Library Screening.** In a typical experiment, 1  $\mu\text{l}$  of a DMSO solution containing a library of 8–15 disulfide-containing compounds is added to 49  $\mu\text{l}$  of protein-containing buffer. These compounds are chosen so that each has a unique molecular weight; ideally, these molecular weights differ by at least 10 atomic mass units so that deconvolution is unambiguous. Although we have typically chosen to screen pools of 8–15 disulfide-containing compounds for ease of deconvolution, larger pools can be used as discussed below and as shown in Fig. 2C. The protein is present at a concentration of  $\approx 15 \mu\text{M}$ , each of the disulfide library members is present at  $\approx 0.2 \text{ mM}$ , and thus the total concentration of all disulfide library members is  $\approx 2 \text{ mM}$ . The reaction is done in a buffer containing 25 mM potassium phosphate (pH 7.5) and 1 mM 2-mercaptoethanol, although other buffers and reducing agents can be used. The reactions are allowed to equilibrate at ambient temperature for at least 30 min. These conditions can be varied considerably depending on the ease with which the protein ionizes in the mass spectrometer (see below), the reactivity of the specific cysteine(s), etc. In the case of TS, the conditions described above were found to be satisfactory. No special effort was made to exclude oxygen or adventitious metal ions; on the time scale of these reactions, there is sufficient free thiol to facilitate disulfide exchange.

After equilibration, the reaction is injected onto an HP1100 HPLC and chromatographed on a  $\text{C}_{18}$  column attached to a mass spectrometer (Finnigan-MAT LCQ, San Jose, CA). The multiply charged ions arising from the protein are deconvoluted with available software (XCALIBUR) to arrive at the mass of the protein. The identity of any library member bonded through a disulfide bond to the protein is then easily determined by subtracting the known mass of the unmodified protein from the observed mass. This process assumes that the attachment of a library member does not dramatically change the ionization characteristics of the protein itself, a conservative assumption because in most cases the protein will be at least 20-fold larger than any given library member. This assumption was confirmed by demonstrating that small molecules selected by one protein are not selected by other proteins (data not shown).

**Crystallography.** Crystals were grown as previously described (15), with the exception that for the noncovalent complexes, 1 mM compound was included in the crystallization buffer. Before data collection, crystals were transferred to a solution containing 70% saturated  $(\text{NH}_4)_2\text{SO}_4$ , 20% glycerol, 50 mM  $\text{K}_2\text{HPO}_4$ , pH 7.0. For the noncovalent *N*-tosyl-D-proline complex, 10 mM compound was added to the soaking solution; for the other complexes, 1 mM compound was included. Diffraction data were collected at  $-170^\circ\text{C}$  by using a Rigaku (Tokyo) RU-3R generator and an *R*-axis-IV detector and processed by using D\*TREK (16). As these crystals were isomorphous with previously described structures [Protein Data Bank code 1TJS (10) for the 12<sub>3</sub> form and 2TSC (17) for the P6<sub>3</sub> form], refinement began by rigid body refinement by using REFMAC [CCP4 (18)]. The protein model was adjusted by using O (19), compound model constructed in INSIGHT-II (Molecular Simulations, Waltham, MA),



**Fig. 2.** (A) Two representative tethering experiments. The protein TS is present at a concentration of 15  $\mu$ M, and each of the 10 disulfide library members in each pool is present at 200  $\mu$ M. The buffer contains 25 mM potassium phosphate (pH 7.5) and 1 mM 2-mercaptoethanol, and the samples were allowed to equilibrate at ambient temperature for 1 h before analysis. (B) Three tethering experiments in which the concentrations of 2-mercaptoethanol were varied as stated. The pool of disulfides is the same as in A Right, and the conditions (other than 2-mercaptoethanol concentration) were the same as above. (C) Three tethering experiments in which the pool size and concentrations were varied as stated. All other conditions were the same as above.

and PROTON [CCP4 (18)] dictionary created by using MAKEDIC [CCP4 (18)]. Positional and individual isotropic temperature factor refinements were carried out with REFMAC [CCP4 (18)] by using all reflections in the indicated resolution ranges. Solvent molecules were placed automatically by using ARPP (CCP4) and refinement continued until no interpretable features remained in  $F_o - F_c$  difference maps. Protein Data Bank accession numbers are 1F4B, 1F4C, 1F4D, 1F4E, 1F4F, and 1F4G for the native C146-tethered *N*-tosyl-D-proline, L143C-tethered *N*-tosyl-D-proline, *N*-tosyl-D-proline free acid soak, glutamate-*N*-tosyl-D-proline soak, and glutamate-*N*-tosyl-D-proline- $\beta$ -alanine crystals, respectively.

## Results and Discussion

The TS protein from *E. coli* contains an active-site cysteine as well as four nonconserved cysteine residues that are buried in the structure (15). Previous reports suggested that the active-site cysteine (C146) is particularly susceptible to modification (20).

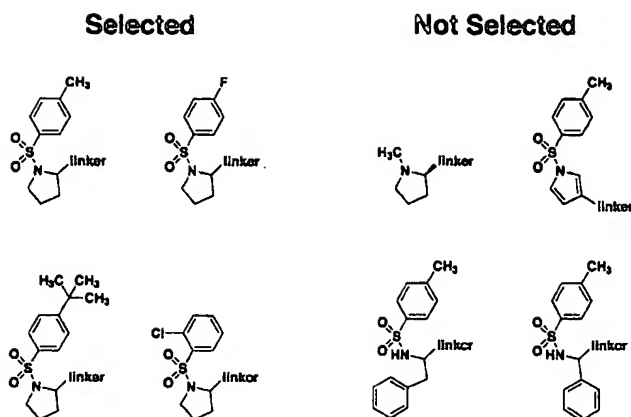
This was confirmed by reacting either wild-type or C146S TS (in which the active-site Cys-146 has been mutated to Ser) with cystamine,  $H_2NCH_2CH_2S-SCH_2CH_2NH_2$ . The results (not shown) revealed that whereas wild-type enzyme cleanly reacted with one equivalent of cystamine, the mutant TS did not react. This selectivity gave us confidence that in our screening experiments the site of modification would be the active-site cysteine.

Two representative tethering experiments are shown in Fig. 2A. On the left side is the deconvoluted mass spectrum of TS reacted with a pool of 10 different disulfide-containing molecules that do not specifically tether to TS. The large peak on the far left of the spectrum is unmodified TS (mass of 30,479). The slightly larger peak is TS disulfide-bonded to 2-aminoethanethiol (combined mass of 30,553). In the absence of any binding interactions, TS is statistically more likely to react with the 2-aminoethanethiol portion of the disulfide moiety than any individual library member because 2-aminoethanethiol is common to all of the library members. The small higher-mass peaks correspond to discrete library members, none of which is prominent. The spectrum on the right side of Fig. 2A shows the result of reacting TS with a different pool of 10 disulfide-containing molecules. In this case, the peaks corresponding to free TS and 2-aminoethanethiol-modified TS are dwarfed by a peak (mass of 30,805) that corresponds to TS disulfide bonded to the *N*-tosyl-D-proline compound shown.

For disulfide tethering to capture the most stable ligand, the reaction must be under rapid exchange to allow for equilibration. Moreover, it should be possible to increase the stringency of the selection to capture more stable ligands by addition of a reducing agent. Both these concepts can be tested by titration with 2-mercaptoethanol (Fig. 2B). In the absence of added reducing agent (far left), *N*-tosyl-D-proline is only moderately selected. This is because the initial reaction of the active-site cysteine with a disulfide is kinetically controlled. Under these conditions, most of the TS is modified with 2-aminoethanethiol. However, on addition of catalytic amounts of reducing agent (Middle), *N*-tosyl-D-proline is strongly selected. Even under strongly reducing conditions (Right), the covalent adduct with *N*-tosyl-D-proline disulfide is still prominent. These results demonstrate that the reaction with *N*-tosyl-D-proline disulfide is a thermodynamic rather than a kinetic process. Thus the selection is based in part on the actual binding energy of the *N*-tosyl-D-proline analog to TS.

To maximize throughput and efficiency, it is desirable to screen multiple compounds simultaneously. Because MS is used to deconvolute the pools, as long as each compound has a unique molecular weight, arbitrarily large pools can be screened. Fig. 2C shows three different tethering experiments in which *N*-tosyl-D-proline was selected from increasingly larger and more diverse pools at decreasing concentrations. As seen in the Fig. 2C Right, *N*-tosyl-D-proline is still readily selected even in the presence of 100 compounds. Each experiment requires only minutes to run, so in principle the technique is amenable to high-throughput screening.

After screening a library of 1,200 compounds, we started to observe some highly significant structure-activity relationships (SAR), as shown in Fig. 3. There is evidently a fair amount of flexibility around the phenyl ring, as shown by the fact that the methyl group can be replaced by a *t*-butyl group or removed entirely. However, the phenyl-sulfonamide core moiety appears to be essential, because methyl proline is not selected. Furthermore, the proline ring itself appears to be quite important; replacing it with a phenylalanine, phenylglycine, or pyrrole causes the resulting compound not to be selected. Of course, these data are qualitative and are not meant to be exhaustive. For example, although the pyrrole compound shown was not selected, this could be because of linker derivitization at the 3-position while proline is substituted at the 2-position. Although it would be possible to greatly

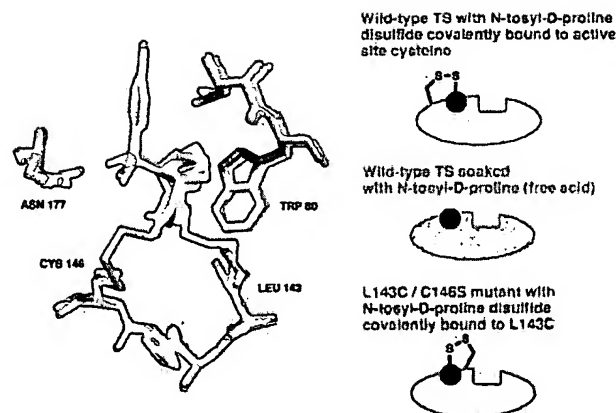


**Fig. 3.** Compounds either selected by covalent tethering (Left) or present in the disulfide library but not selected (Right). Except in the case of *N*-tosyl-D-proline, all compounds tested were racemic unless otherwise indicated. For *N*-tosyl-D-proline, both stereoisomers were screened separately in different pools, and both were identified as hits, although the D-isomer appeared to be selected slightly more strongly than the L-isomer.

expand the disulfide library to try to obtain detailed SAR, we felt it would be more efficient to remove the disulfide as soon as possible to determine the actual binding affinity of the untethered pharmacophore.

We determined the affinity of the tether-free analog, *N*-tosyl-D-proline, by Michaelis–Menten analysis. Enzymatic assays (12, 13) were performed to determine the inhibition constant ( $K_i$ ). The  $K_i$  for the free acid was found to be  $1.1 \pm 0.25$  mM and was competitive with respect to the substrate dUMP. Thus, *N*-tosyl-D-proline is a weak but competitive inhibitor of TS.

We next evaluated the set of ligands that could be captured by a nearby cysteine. To do this, we removed the active-site cysteine (C146S) and introduced a new cysteine nearby (L143C or H147C). Screening the disulfide-containing library against the C146S/L143C mutant TS produced similar results to the wild-type enzyme: the *N*-tosyl-D-proline analog was strongly selected (data not shown). In contrast, screening the disulfide-containing library



**Fig. 4.** Overlay of three crystallographically determined structures. The structure in green was determined after soaking *N*-tosyl-D-proline (free acid) into crystals of unmodified TS. The structure in red is TS covalently modified by *N*-tosyl-D-proline disulfide-bonded to C146 (the active-site cysteine). Finally, the structure in blue is mutant TS (C146S/L143C) covalently modified by *N*-tosyl-D-proline disulfide bonded to L143C.

against the C146S/H147C mutant TS did not result in selection of *N*-tosyl-D-proline. Several other molecules were selected by these mutants, but none were as prominent as *N*-tosyl-D-proline. These data demonstrate that the selection of *N*-tosyl-D-proline is not unique to the active-site cysteine, and that the exact placement of the reactive cysteine residue is not critical. As expected, however, there is some dependence on the position of the cysteine.

To determine the mode of tosyl-D-proline binding, either free or when tethered from C146 or L143C, we solved the x-ray crystallographic structures of the complexes (Table 1 and Fig. 4). Significantly, the location of the *N*-tosyl-D-proline moiety is very similar in all three cases (rms deviation of 0.55–1.88 Å, compared with 0.11–0.56 Å for all  $C_\alpha$  carbons in the protein). The fact that the *N*-tosyl-D-proline substituents closely overlap while the alkyl-disulfide tethers converge onto this moiety from different cysteine residues supports the notion that the *N*-tosyl-D-proline moiety, not the tether, determines binding.

These results demonstrate that the tethering methodology can rapidly identify small low-affinity ligands. However, to be useful

**Table 1.** Crystallographic data and refinement parameters

| Data set                  | Space group*      | Cell dimensions, Å   | Resolution, Å | Reflections (no.) |        | Completeness, <sup>†</sup> % | $R_{\text{sym}}$ ( $I$ ), <sup>‡</sup> % | $I/\sigma$ | $R_{\text{cryst}}$ , <sup>§</sup> % | $R_{\text{free}}$ , <sup>  </sup> % | rms deviation bond lengths, Å | rms deviation bond angles, deg |
|---------------------------|-------------------|----------------------|---------------|-------------------|--------|------------------------------|--|------------|-------------------------------------|-------------------------------------|-------------------------------|--------------------------------|
|                           |                   |                      |               | Overall           | Unique |                              |  |            |                                     |                                     |                               |                                |
| Native                    | I2 <sub>1</sub> 3 | a = 131.17           | 10 – 1.75     | 104,019           | 36,586 | 96.7 (91.6)                  | 4.9 (33.8)                               | 20.5 (4.0) | 19.8                                | 24.4                                | 0.010                         | 2.30                           |
| C146 tethered             | P6 <sub>3</sub>   | a = 126.22 c = 67.02 | 10 – 2.00     | 97,445            | 41,001 | 98.8 (94.5)                  | 4.4 (26.0)                               | 14.7 (4.1) | 19.8                                | 26.8                                | 0.010                         | 2.59                           |
| <i>N</i> -tosyl-D-proline |                   |                      |               |                   |        |                              |  |            |                                     |                                     |                               |                                |
| L143C tethered            | P6 <sub>3</sub>   | a = 126.33 c = 67.12 | 10 – 2.15     | 78,793            | 32,045 | 96.7 (92.1)                  | 8.1 (28.6)                               | 12.8 (4.5) | 19.6                                | 26.7                                | 0.014                         | 3.06                           |
| <i>N</i> -tosyl-D-proline |                   |                      |               |                   |        |                              |  |            |                                     |                                     |                               |                                |
| Noncovalent               | I2 <sub>1</sub> 3 | a = 131.88           | 10 – 1.90     | 202,300           | 31,422 | 100 (100)                    | 7.4 (28.2)                               | 19.7 (3.8) | 19.2                                | 23.8                                | 0.011                         | 2.49                           |
| <i>N</i> -tosyl-D-proline |                   |                      |               |                   |        |                              |  |            |                                     |                                     |                               |                                |
| Glu-TP                    | P6 <sub>3</sub>   | a = 126.14 c = 66.81 | 10 – 2.00     | 143,599           | 40,497 | 99.4 (96.9)                  | 8.5 (31.9)                               | 13.9 (4.0) | 19.4                                | 25.1                                | 0.007                         | 2.15                           |
| Glu-TP-β-Ala              | P6 <sub>3</sub>   | a = 126.03 c = 66.84 | 10 – 1.75     | 142,016           | 58,487 | 95.8 (85.2)                  | 4.0 (22.5)                               | 17.1 (4.9) | 18.0                                | 21.4                                | 0.007                         | 2.00                           |

This is not a “true” free  $R$  factor because the starting model was a fully refined structure. However, the free  $R$  factor set of reflections was kept constant for each of the above refinements.

\*The I2<sub>1</sub>3 crystal contains one monomer per asymmetric unit. The P6<sub>3</sub> form contains the biologically relevant homodimer.

<sup>†</sup>Values in parentheses are for the highest resolution bin.

<sup>‡</sup> $R_{\text{sym}} (I) = \sum h_{\text{hkl}} / \sum (h_{\text{hkl}} / I_{\text{hkl}})$ , where  $I_{\text{hkl}}$  is the intensity of reflection  $h_{\text{hkl}}$ .

<sup>§</sup> $R_{\text{cryst}} = \sum |F_{\text{obs}}| - |F_{\text{calc}}| / \sum |F_{\text{obs}}|$ , where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure factors, respectively, for the data used in refinement.

<sup>||</sup> $R_{\text{free}} = \sum |F_{\text{obs}}| - |F_{\text{calc}}| / \sum |F_{\text{obs}}|$ , where  $F_{\text{obs}}$  and  $F_{\text{calc}}$  are the observed and calculated structure factors, respectively, for 10% of the data omitted from refinement.

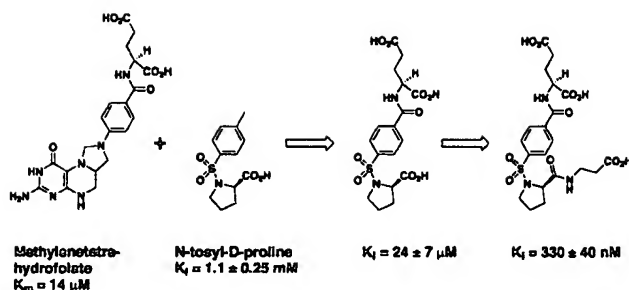


Fig. 5. Grafting a glutamate residue onto *N*-tosyl-D-proline improves the affinity 50-fold, and adding a negatively charged appendage further increases the affinity by an additional 70-fold.

for drug discovery, the ligands identified must be amenable to medicinal chemistry and affinity maturation. In the case of *N*-tosyl-D-proline, examination of the bound structures revealed that the tosyl group is in roughly the same position and orientation as the benzamide moiety of methylenetetrahydrofolate, the natural cofactor for this enzyme (17, 21, 22). Given this positioning, we grafted the glutamate residue from methylenetetrahydrofolate onto the methyl group of *N*-tosyl-D-proline to try to enhance the affinity of the molecule (Fig. 5).

A series of compounds was synthesized and tested for TS inhibition. The parent compound, in which L-glutamate is grafted onto *N*-tosyl-D-proline, is almost 50-fold more active than *N*-tosyl-D-proline alone ( $K_i = 24 \pm 7 \mu\text{M}$ ). The  $\alpha$ -carboxylate of the glutamate residue is very sensitive to modification, in that converting it to a primary amide reduces the affinity by an order of magnitude (not shown). Several substitutions were also incorporated at the proline carboxyl group; in one case, the displacement of the negatively charged carboxyl group away from the proline ring by two methylene units improves the inhibition constant by more than 70-fold, to  $K_i = 330 \pm 40 \text{ nM}$  (Fig. 5). Converting this  $\beta$ -alanine appendage to the isosteric isomyl group [i.e., converting  $-\text{NHCH}_2\text{CH}_2\text{CO}_2\text{H}$  to  $-\text{NHCH}_2\text{CH}_2\text{CH}(\text{CH}_3)_2$ , not shown] decreases the binding affinity by more than an order of magnitude ( $K_i = 12 \pm 2.5 \mu\text{M}$ ), suggesting that the displaced carboxylate is critical for improved binding. The crystallographic structures of TS complexed with two of the improved inhibitors (Table 1 and Fig. 6) reveal that the proline ring and the tosyl group overlap in all three structures, whereas the  $\beta$ -alanine and glutamate appendages make contacts with surrounding protein residues. The combined effects of these added substituents increase the affinity of one of the inhibitors for TS by more than 3,000-fold over *N*-tosyl-D-proline. The affinity of this inhibitor is submicromolar and thus well within the range of typical drug leads.

In this study, we have focused on disulfide tethers. Although it is possible to use other types of reversible bonds, such as imines (23), the disulfide bond is well suited for tethering. Disulfide bonds can be formed and broken under very mild conditions with complete chemoselectivity. Moreover, the disulfide moiety itself can be easily

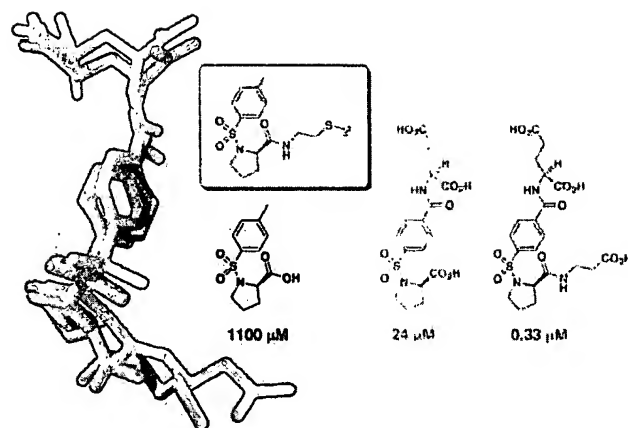


Fig. 6. Overlay of three crystallographically determined structures; only the inhibitor is shown for clarity. The inhibition constant ( $K_i$ ) of each inhibitor is also shown.

introduced into a wide variety of small molecules at the end of a flexible and variable length tether. Finally, previous research has demonstrated that the redox potential of a disulfide bond in a macromolecule can vary with the conformation of the molecule, and that an intermolecular disulfide bond varies in stability as a function of the affinity of the two macromolecules (24–26). Thus, disulfide stability correlates with binding affinity.

A reasonable extension of the tethering technology would be to discover two weakly binding fragments that bind near one another and to link them to produce higher affinity compounds (27–31). Such a strategy has been applied in the technique “SAR by NMR” (32). In many cases, the new linked compounds bind to the target protein with much higher affinity than the precursors (33). Our tethering approach can also rapidly generate candidate molecules for linking.

In conclusion, we have developed an alternative and general approach for drug design that can be readily supplemented with existing technology. The approach allows us to identify ligands that bind very weakly to proteins and thus expands the discovery range for small molecule drugs. The weakly binding ligands can be rapidly optimized by using medicinal chemistry and structure-aided design. This approach is well suited to proteins containing a cysteine in the active site and is also expandable to proteins in which cysteines are introduced by mutagenesis. In principle, new cysteines can be placed anywhere; for example, if a cysteine is introduced onto the surface of a protein in an area known to be important for protein–protein interactions, small molecules could be selected that bind to and block this surface. We expect this covalent tethering methodology to be a powerful technique for generating drug leads.

We thank Dennis Lee and Jon Ellman for helpful discussions and Monya L. Baker for a critical reading of this manuscript. This research was supported in part by Small Business Innovation Research Grant No. 1 R43 CA85141–01.

- Santi, D. V. & Danenberg, P. V. (1984) in *Chemistry and Biology of Folate*, eds. Blakley, R. L. & Benkovic, S. J. (Wiley, New York), Vol. 1, pp. 345–398.
- Carreras, C. W. & Santi, D. V. (1995) *Annu. Rev. Biochem.* **64**, 721–762.
- Kamen, B. (1997) *Semin. Oncol.* **5** Suppl 18, S30–S39.
- Jones, T. R., Calvert, A. H., Jackman, A. L., Brown, S. J., Jones, M. & Harrap, K. R. (1981) *Eur. J. Cancer* **17**, 11–19.
- Ivanetich, K. M. & Santi, D. V. (1990) *FASEB J.* **4**, 1591–1597.
- Vasquez, J. R., Gooze, L., Kim, K., Gut, J., Petersen, C. & Nelson, R. G. (1996) *Mol. Biochem. Parasitol.* **79**, 153–165.
- Webber, S. E., Bleckman, T. M., Attard, J., Deal, J. G., Kathardecarr, V., Welsh, K. M., Webber, S., Janson, C. A., Matthews, D. A., Smith, W. W., et al. (1993) *J. Med. Chem.* **36**, 733–746.

- Shoichet, B. K., Stroud, R. M., Santi, D. V., Kuntz, I. D. & Perry, K. M. (1993) *Science* **259**, 1445–1450.
- Stout, T. J., Tondi, D., Rinaldi, M., Barlocco, D., Pecorari, P., Santi, D. V., Kuntz, I. D., Stroud, R. M., Shoichet, B. K. & Costi, M. P. (1999) *Biochemistry* **38**, 1607–1617.
- Stout, T. J., Sage, C. R. & Stroud, R. M. (1998) *Structure (London)* **6**, 839–848.
- Climie, S. C., Carreras, C. W. & Santi, D. V. (1992) *Biochemistry* **31**, 6032–6038.
- Chen, S. C., Daron, H. H. & Aull, J. L. (1989) *Int. J. Biochem.* **21**, 1217–1221.
- Wahba, A. L. & Friedkin, M. (1961) *J. Biol. Chem.* **263**, PC11–PC12.
- Parlow, J. J. & Normansell, J. E. (1995) *Mol. Diversity* **1**, 266–269.

15. Perry, K. M., Fauman, E. B., Finer-Moore, J. S., Montfort, W. R., Maley, G. F., Maley, F. & Stroud, R. M. (1990) *Proteins* **8**, 315–333.
16. Pflugrath, J. W. (1999) *Acta Crystallogr. D* **55**, 1718–1725.
17. Montfort, W. R., Perry, K. M., Fauman, E. B., Finer-Moore, J. S., Maley, G. F., Hardy, L., Maley, F. & Stroud, R. M. (1990) *Biochemistry* **29**, 6964–6977.
18. Collaborative Computational Project, N (1994) *Acta Crystallogr. D* **50**, 760–763.
19. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47**, 110–119.
20. Aull, J. L. & Daron, H. H. (1980) *Biochim. Biophys. Acta* **614**, 31–39.
21. Matthews, D. A., Appelt, K., Oatley, S. J. & Xuong, N. H. (1990) *J. Mol. Biol.* **214**, 923–936.
22. Kamb, A., Finer-Moore, J. S. & Stroud, R. M. (1992) *Biochemistry* **31**, 12876–12884.
23. Huc, I. & Lehn, J.-M. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 2106–2110.
24. Gilbert, H. F. (1995) *Methods Enzymol.* **251**, 8–28.
25. Stanojevic, D. & Verdine, G. L. (1995) *Nat. Struct. Biol.* **2**, 450–457.
26. O'Shea, E. K., Rutkowski, R., Stafford, W. F., III & Kim, P. S. (1989) *Science* **245**, 646–648.
27. Jencks, W. P. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4046–4050.
28. Portoghese, P. S., Ronsisvalle, G., Larson, D. L. & Takemori, A. E. (1986) *J. Med. Chem.* **29**, 1650–1653.
29. Glick, G. D., Toogood, P. L., Wiley, D. C., Skehel, J. J. & Knowles, J. R. (1991) *J. Biol. Chem.* **266**, 23660–23669.
30. Wetterau, J. R., Gregg, R. E., Harrity, T. W., Arbeen, C., Cap, M., Connolly, F., Chu, C.-H., George, R. J., Gordon, D. A., Jamil, H., *et al.* (1998) *Science* **282**, 751–754.
31. Maly, D. J., Choong, I. C. & Ellman, J. A. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 2419–2424.
32. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. (1996) *Science* **274**, 1531–1534.
33. Hajduk, P. J., Sheppard, G., Nettesheim, D. G., Olejniczak, E. T., Shuker, S. B., Meadows, R. P., Steinman, D. H., Carrera, G. M., Marcotte, P. A., Severin, J., *et al.* (1997) *J. Am. Chem. Soc.* **119**, 5818–5827.